APPENDIX I

Methodology for 3-Mile Radius Demographic Profiles

This appendix describes the methods used to compile two types of data for areas surrounding assembly plants: (1) demographic data taken from the Census of Population for various sized areas around plants (Census Blocks and circles of 0-1, 0-3, 1-3, and 3-5 mile radii around plants), and (2) 1994 TRI emissions data for sources located within a three-mile radius. Assembly plant locations were identified by the lat/long coordinates described in Appendix F. Geographic Information System (GIS) software was used by EPA Region IV staff to calculate sums for the areas of interest, as described below.

Demographic Data

All demographic data used were obtained from the 1990 U.S. Census of Population Summary Tape File version 3A. The Block Group summary level data were used to calculate populations in the vicinity of each assembly plant site. Summary levels 040 and 050 were used to extract comparative data for the states and counties. It was assumed that Census Block populations and their racial, income and educational attainment characteristics are uniformly distributed within the Census Block, which allowed estimating numbers of persons in various categories for pieces of Census Blocks that fall within the target circles around plants

The following are the sources and methods used for each data element:

- Total Persons was taken from STF Table P0010001

- Minority Population was calculated by subtracting the reported White Persons (STF Table P008001) from STF Table P0010001 (Total Persons).

- Educational Attainment for persons aged 18 and older is reported as percentage not completing high school. This percentage was calculated by summing the values for STF Tables P0600001 and P0600002, which represent, respectively, a count of persons who have less than a 9th grade education and a count of persons having achieved an education between 9th and 12th grades, inclusive, but not obtaining a diploma. Total educated persons aged 18 or older was calculated by summing STF Tables P0600001 through P0600007. The percentage was then calculated as:

$$100 * \frac{\{\text{\# educated persons without high school diploma}\}}{}$$

{Total # educated persons}

- Poverty Level was calculated using the national poverty level income of $12,500 per four-person household. A household consists of either one person living alone or any group of related or unrelated persons living together in a single housing unit. The total number of households was calculated by summing STF Tables P0800001 through P0800025, and the number of households below the four-person household poverty level was calculated by summing Tables P0800001, P0800002 and P0800003. The percent of households below the poverty level was then calculated as:

$$100 * \frac{\{\text{\# households below poverty level income}\}}{\{\text{Total \# households}\}}$$

  Data for county-level poverty rates differ as shown on page II-29 (taken directly from Census publications) and on page II-34 and the Part III profiles, due to differences in the methodologies used. The method used to calculate the county-level data shown on page II-29 is more precise, but could not be used for areas below the county level. A more precise estimate of percentages of households below the poverty level would compare income for each household to the appropriate poverty level income for that size household. This more accurate calculation could not be performed because data on household income are not reported in sufficient detail for reasons of confidentiality. The approximation used may either over- or understate numbers of households below the poverty level, depending on the size distribution of households in a particular area. To provide a consistent comparison across geographic units, the county- and state-level data shown on page II-34 and the Part III profiles use the less precise method described here.

- Population densities were calculated based on land area only.

Toxics Release Inventory Data

To identify TRI reporters within three miles of each assembly plant, the facility lat/long was "buffered" with a three mile radius. This buffer was then used to clip from a TRI dataset the sites within three miles. The TRI coverage was obtained from the national TRI coverages available on the EPA internet web site.

TRI release data were downloaded from EPA's mainframe using a standard retrieval program.

The fields obtained included TRI ID number, CAS codes, Chemical Name, and Release Type Codes (M codes). The data were then processed to ensure that only one CAS number/TRI number pair existed for each chemical released from a particular facility. The fields were then filled in by inserting values using the original chemical table and a subset table in a relational schema.